Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

MEASUREMENT OF GROWTH AND EFFICIENCY IN ARITHMETIC (*Concluded*)

S. A. COURTIS
Home and Day School, Detroit

Of equal importance with the question of the amount of growth produced by a year's work is that of the permanence of the growth. Table VI gives for Grades 4 to 9 the average scores in the speed tests in September, 1909, the highest scores made at any time during the year, and the scores of the same classes the following September after the long summer vacation. From these the gains during the year, the losses during the summer can easily be computed. The last two columns show the differences between the September tests and the percentage of improvement.

The results are striking and suggestive. The large gains of the fifth and sixth grades are offset by the large losses during vacation, while the growth of the seventh and eighth grades is seen to be relatively permanent. If these results are confirmed by future work, the seventh and eighth grades would seem to be the place in the course where rigid drill on the combinations would be most effective.

At the beginning of the year it was felt by all concerned that there was one question that the year's work must settle, and settle very definitely. Do such short, simple, comparative tests yield reliable results, or would repeated tests give very different relative standings? There is not space here to marshal all the evidence that proves the grade averages reliable measures of grade abilities, but an examination of the results in Table V[1] from this point of view will show how consistently grade after grade has obtained the same score in successive tests. It is to be remembered that all these grades were subject to constant practice in their daily work and drill, so that it is only the before-and-after-vacation scores, or successive scores during a

---

[1] See *Elementary School Teacher* (March, 1911), 367.

"plateau" that can be expected to be constant. Such scores are, however, remarkably constant. Extended practice, as has been shown, does produce a rise in score, yet but a part of this is permanent, and the abilities in question, slowly built up by hours of study and drill, must be considered relatively stable quantities. The constancy of very many grade and individual scores in repeated tests proves conclusively that the results obtained from even a single test carefully conducted are reliable measures of grade ability.

TABLE VI

GAINS AND LOSSES

GRADE	OPERATION	SCORE		MAXIMUM GAIN	SCORE	SEPTEMBER 1910 (Loss during Summer Vacation)	GAIN 1910 OVER 1909	PERCENTAGE
		September 1909	Highest		September 1910			
4	+	23	39	16	37	2	14	
	−	19	37	16	33	4	14	
	×	6	22	16	17	5	11	
	÷	7	25	18	17	8	10	
	Average	13.7	..	16.5	..	4.7	12.2	89
5	+	35	66	31	47	19	12	
	−	30	54	24	46	18	16	
	×	24	56	32	34	22	10	
	÷	19	55	36	31	24	12	
	Average	27	..	30.7	..	20.7	12.5	46
6	+	43	73	30	56	17	13	
	−	33	58	25	54	5	20	
	×	30	58	28	42	16	12	
	÷	26	56	30	44	12	18	
	Average	33	..	30.7	..	12.5	14.7	44
7	+	53	89	36	79	10	26	
	−	36	69	33	59	10	23	
	×	41	77	36	64	13	23	
	÷	34	74	40	62	12	28	
	Average	41	..	36.2	..	11.2	25	61
8	+	56	85	29	79	6	23	
	−	42	70	28	64	6	22	
	×	49	74	25	60	14	21	
	÷	41	77	36	67	10	26	
	Average	47	..	29.5	..	9	23	49
9	+	67	81	24	71	10	4	
	−	45	64	19	63	1	18	
	×	56	69	13	68	1	12	
	÷	46	64	18	61	3	15	
	Average	53.5	..	18.5	..	4	12.2	22

On the other hand it is important to state that this is not true of the scores of the individuals composing the grade. Even when the grade average is perfectly constant, individual scores are found to vary. This is shown clearly in Table VII, the individual records in the addition speed test of 23 of the 31 members of the eighth grade. The records omitted are those in which through absence the series was not complete. The graph (Fig. 12) makes possible the rapid visual comparison of the changes in the individual scores with those of the grade average, the heavy central line. It will be seen that at such times as the grade average is constant, the constancy is due to the fact that the variations of certain individuals in one direction are offset by the variations of others in the opposite direction. For the individual, therefore, repeated tests are needed to determine his most frequent score and the range of his variation from it.

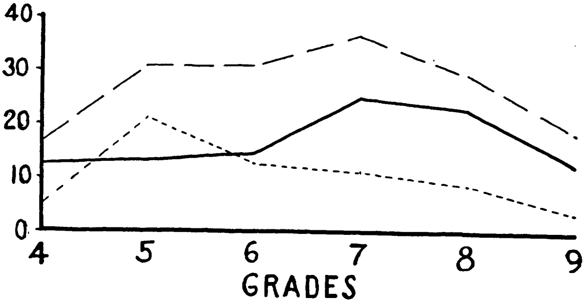

FIG. 11.—Gains and losses for the year. Broken line shows maximum gain—from September, 1909, to highest point reached at any time during year. Dotted line shows effect of summer vacation—loss from highest point reached to September, 1910. Full line shows net gain for year, September, 1910, over September, 1909.

By far the most important fact to be gained from a study of Table VII is the wide range of ability within the grade. The score of the best girl is on the average twice that of the poorest girl, in spite of the fact that the individuals in this class represent the survivals of eight years of "passing" on a 70 per cent basis. The graph shows plainly that the grade average is really representative of very few scores, and that a measure of the range of variation is needed in addition to the measure of the average performance. The writer has chosen the average devia-

TABLE VII

SCORES IN ADDITION SPEED TEST OF TWENTY-THREE INDIVIDUALS OF EIGHTH GRADE

Rank	September 1909	November	December	January	March 1–18	March 18–30	April				May					June	September 1910	Rank
							4	11	18	25	2	9	16	23	30	6		
1	89	85	..	..	86	91	97	98	98	102	..	86	..	105	120	100	88	7
2	77	78	80	70	78	57	83	81	77	82	..	85	..	84	91	83	..	10
3	69	81	87	91	90	107	107	116	111	..	..	98	..	115	120	109	104	1
4	68	81	83	85	90	91	93	94	101	107	..	97	..	104	116	99	92	4
5	65	82	77	79	83	84	85	91	85	98	..	90	..	104	100	97	..	5
6	64	83	87	94	84	100	100	110	105	107	..	100	..	108	120	106	97	2
7	62	55	60	58	64	..	66	74	70	70	..	73	..	73	77	79	74	14
8	58	63	68	64	77	72	80	80	73	84	..	80	..	80	88	83	80	11
9	57	77	72	76	75	79	83	80	84	90	..	81	..	90	96	86	83	8
10	56	77	80	78	80	76	86	85	97	102	..	94	..	105	110	88	90	6
11	55	69	75	74	77	82	83	84	87	87	..	85	..	84	83	90	83	9
12	52	69	71	77	84	86	90	99	90	94	..	90	..	92	92	100	97	3
13	52	60	51	57	..	80	83	80	82	86	..	79	..	80	70	81	72	16
14	52	53	75	50	68	60	78	80	82	71	..	66	..	67	77	80	76	12
15	51	65	55	62	68	63	73	69	76	77	..	78	..	83	88	82	..	17
16	46	51	54	47	57	59	55	64	60	66	..	69	..	62	73	69	..	18
17	46	40	39	43	48	47	55	54	60	61	..	62	..	66	71	67	..	19
18	44	65	..	60	68	..	72	71	80	80	..	80	..	80	78	83	73	15
19	44	58	68	70	67	72	77	..	75	74	..	82	..	75	85	80	76	13
20	43	30	50	56	60	59	59	61	55	61	..	58	..	55	52	64	59	20
21	35	43	50	27	52	40	53	51	42	..	..	42	..	55	57	51	48	23
22	31	30	..	39	42	..	49	57	61	55	..	52	..	54	50	53	49	22
23	30	40	45	47	48	58	55	54	50	57	..	57	..	56	55	59	..	21
Aver.	54	62	66	64	70	73	77	79	79	81	..	78	..	82	85	82	79	..
A.D.	10.8	14.3	12.2	14.2	11.9	14.5	13.2	14.1	14.3	13.6	..	12.2	..	15.8	17.3	12	12.3	.
Var.	20	23	19	22	17	20	17	18	18	18	..	16	..	19	22	15	16	..

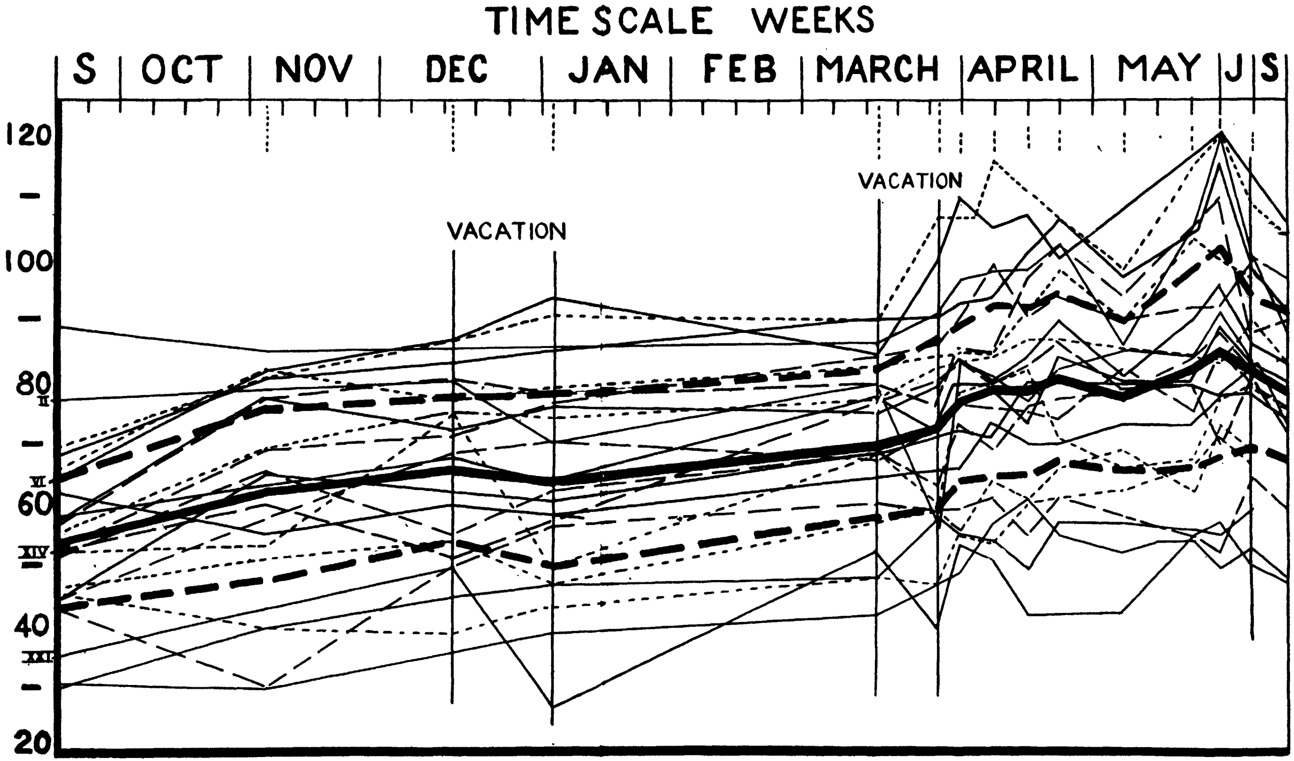

FIG. 12.—Records of twenty-three individuals of the eighth grade in fifteen trials of the speed test in addition from September, 1909, to September, 1910. Full, heavy line shows the average score. Heavy broken lines are at a distance from the average equal to the average deviation. The zone bounded by the two broken lines corresponds to a variability of 18 per cent.

tion, the average difference of the individual score from the grade average, as the measure of variability, and for purposes of comparison the average deviations will be expressed as percentages of their respective averages. The significance of the measure is shown graphically in the plot. On each side of the heavy line representing the grade average, there has been drawn a heavy broken line at a distance from the average equal to that of the average deviation. Between the two broken lines will be found roughly 50 per cent of the class. That is, the exceptionally weak and the exceptionally able have been excluded, and the upper and lower limits of ability of the remainder of the class are given by the broken lines. Because of this wide individual range, therefore, the ability of a class is better thought of as a "zone" than as a single average score. A narrow zone indicates a compact well-graded class, the product of an efficient system of promotion. Unfortunately the zones from the present tests are wide, wider at least than would seem to the writer to be expected in any school that prides itself on its standards, and he would hesitate to publish such results did not returns from other schools using the same tests under the same conditions show zones of the same or greater width. It seems probable that the condition is general in education at the present time. Within the school other tests and measurements make comparisons possible. The variability of the same 23 individual girls whose scores in addition were given in Table VII, and whose variability was there shown to be 18 per cent, was 3 per cent in age, 7 per cent in handwriting (samples measured by Thorndike's scales), 15 per cent in formal English grammar, 26 per cent in height, 30 per cent in memory of important dates and men, 60 per cent in ability to reproduce after 24 hours the main points of a historical passage read once in class.

No one result of the testing work has proved more disturbing to the writer, more startling in its implications, than this extreme variability found in all the grades. Its full significance can be appreciated only by one who has made a careful study of the individual records and then noted the behavior of the same individuals in the classroom. The writer believes weakness and inefficiency are due more to failure to secure complete, all-around development in the fundamental abilities of a subject, and standard growths in each grade, than to any other one cause. Just how inefficient the present conditions are is shown in the case of knowledge of the multiplication tables by the results in Table VIII and Fig. 13. The table gives the average and the average deviation for each of the four grades in all the multiplication tests of the year. The graph shows the zone of ability of each grade shaded in such a way that ground common to two or more grades can be detected at once. The overlapping of grade upon grade is such that for two-thirds of the year a part of the territory is common to *all four grades,* while a large part of it is at all times common to three grades. The table and figure together make very plain the complete breakdown of the *grade* idea. For it is to be remembered that the zones do not include the extremes, but the central core of the class only. In other words, the variations of individual ability within the grade are far greater than the increase of knowledge or ability from grade to grade. It will be seen, therefore, that just so far as these conditions are truly general, the problem of the degree of variability that can be permitted within a grade is a vital one for graded systems.

TABLE VIII

Comparative Record of Grade Ability in Multiplication Speed Tests throughout the Year

		Sept. 1909	Nov.	Dec.	Jan.	Mar.	Mar.	April				May					June 6	Sept. 1910
								4	11	18	25	2	9	16	23	30		
5	Av.	24	32	39	40	45	46	43	52	50	52	..	52	..	55	52	55	34
	A.D.	8	9	10	11	12	12	12	13	12	17	..	14	..	12	11	15	9
	Var.	35	29	26	28	26	25	28	25	25	33	..	27	..	21	21	28	27
6	Av.	30	34	49	41	49	50	55	56	..	..	58	52	54	56	56	47	42
	A.D.	10	9	12	15	11	14	15	13	..	..	13	12	17	14	14	14	13
	Var.	35	27	25	37	22	28	28	23	..	..	22	24	31	24	24	29	31
7	Av.	41	43	48	48	62	59	59	..	72	70	68	74	72	77	77	74	64
	A.D.	12	12	11	12	14	14	14	..	15	17	14	14	11	12	16	11	15
	Var.	29	28	22	25	22	23	23	..	20	24	21	18	16	16	21	15	23
8	Av.	49	55	60	56	62	68	67	68	..	..	71	..	74	73	72	72	60
	A.D.	11	9	11	12	10	12	11	12	..	..	12	..	15	10	14	12	13
	Var.	22	18	18	21	17	18	17	18	..	..	18	..	20	13	19	17	21

The cause of such breakdown of grade standards is to be

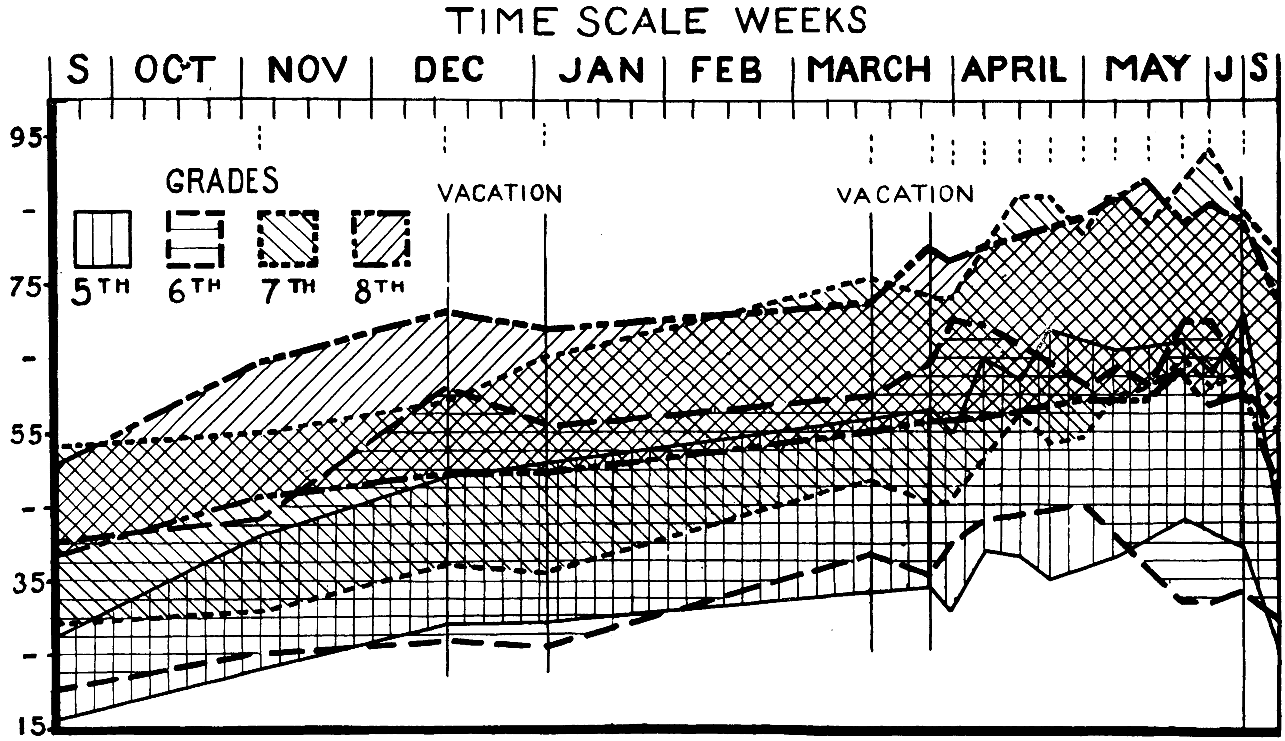

FIG. 13.—Record of year's tests in the multiplication tables for Grades 5–8, showing the extent to which the abilities of the grades overlap. The "zone" of ability for each grade is determined by the grade average and the average deviation from it. Within a zone, therefore, are included merely the abilities of the central core of a class; roughly, the half of the class nearest the grade average.

found in the inherent differences in individual children. Consider the effect of vacation, for instance. In Fig. 12 it will be seen that as a whole the grade score fell slightly during the Christmas vacation and the variability increased, while the opposite was true of the spring vacation. Individuals 2, 14, and 21, however, are consistently demoralized by both vacations. Individual 6, on the other hand, shows just the opposite effect, while some of the others are very slightly affected if at all. One looks in vain for *any* constant effect during the year. The test the last week in May is particularly interesting as bearing on this point. It represents the trial of a new test, similar to the old but better balanced as to units. To many children it proved much easier, but there were individuals in every grade to whom it seemed more difficult, and in the case of the eighth grade the variability was correspondingly increased. From the point of view of the year's growth the same thing is evident. The grade work that enabled No. 12 in rank in September, 1909, to become No. 4 in September, 1910, and No. 18 No. 12, changed the relative rank of other members of the grade very little, while there are a few whose absolute score also shows but little improvement. In other words, education is an individual matter, and no grade instruction nor group work can possibly be devised to minister adequately and efficiently to the widely differing needs of any group selected under the present system.

The writer, however, is not one of those who believe that the needs of the individual will bring about the abolition of our present graded system. The necessity for handling large numbers of children and the many benefits that accrue from group work make its continuance a certainty. At the same time the differences in individuals make it imperative that each child receive the kind and amount of instruction that his particular mental makeup demands. These two conditions can be met by having for each subject and each grade standard scores and standard growths in all those component abilities that are fundamental. Individual measurement by standard tests and constant checking of the results of classwork, supplemented by individual prescriptions of study and work, would insure the attainment of any desired degree of minimum ability, and still leave room in the nonessentials of the subject for all those permissible individual variations which make for personality. If education is to become either scientific or efficient, more attention must be given to accurate determinations of both the material to be acted upon, and the effects produced. In the future it may well be that a testing department will be a vital part of every well-organized school system.

One other question will briefly be discussed, the correlation between the elemental abilities and the complexes of which they form a part. The question of how far drill on the multiplication tables, for instance, must be carried in order that a child may be able to multiply correctly in the solution of problems is one of some importance. Although in all discussions based upon the present tests the defects of the tests themselves and the consequent possible unreliability of the results must be kept in mind, two correlations attempted seem to yield results of value.

The first is the relation between rate of motor activity and total score in the speed tests. It seemed reasonable to suppose that a child able to copy figures more rapidly than another would obtain higher scores in writing the combination, even though the two children really had equal knowledge of the tables. To test this point the June scores of 182 individuals of all grades above the third were thrown into one large group and distributed on the basis of their scores in copying figures. The average total score in each of these smaller groups was then determined. Figure 14 shows the results graphically. It is evident from the curve that there is a very close connection between the two types of scores up to a certain point, but that beyond this point the ability to copy figures with extreme rapidity does not carry with it the corresponding ability to obtain a high score in the other speed tests. This critical score, therefore, approximately 115 figures in a minute, has been taken by the writer as the standard to be reached before the work in arithmetic is completed. Standard scores, graded down from this maximum, have been set for the other classes below the eighth, and the attempt is being made by classwork and by

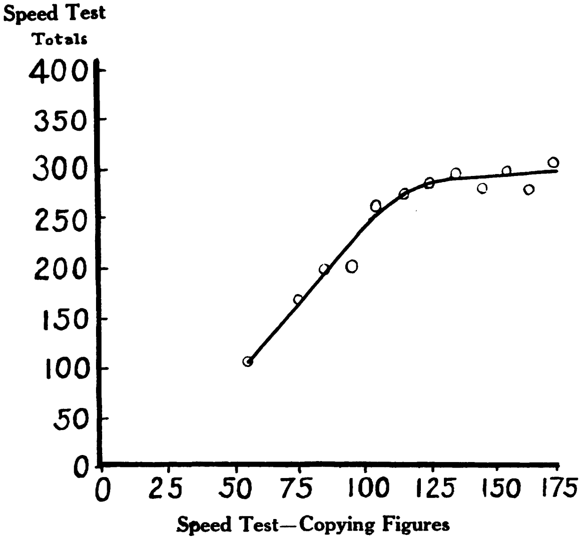

Fig. 14.—Correlation between speed in copying figures and total scores in the speed tests.

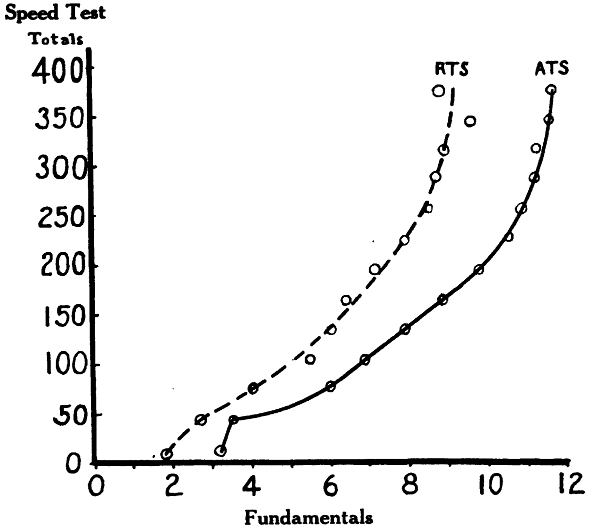

Fig. 15.—Correlation between total scores in the speed tests and the number of examples attempted and right in the test on fundamentals.

extra special individual work to bring the score of every individual up to the score for her grade. By thus securing a standard growth in each grade, the final maximum ability will be insured.

A second correlation of some importance is the relation between total scores in the speed tests and the scores in the tests on abstract examples. Figure 15 gives the results both for examples attempted and for examples right. The general form of the curves is the same. That is, there is probably a critical degree of knowledge of the tables to be attained by drill beyond which an increase in score does not carry with it an increase in ability to work examples. The determination of the proper balance between these fundamental abilities of a subject affords a large number of important problems for experimental work.

No general summary of the points made in the foregoing discussions will be attempted. The writer feels, however, that the work of the year has proved conclusively the necessity of measurements of individual needs and individual growth, as well as the great possibilities of comparative tests from the standpoint of intelligent determination of courses and methods. It has also made possible the construction of tests composed of units of known values, so that future measurements will be free from this source of error. The immediate work of the future is, therefore, the determination of standards for each grade and of standard growths. It must also settle the question of whether or not the teacher will be able, by use of the knowledge gained from the tests, to bring individuals up to standard scores. It is with these questions that the experimental work of the school during the coming year is to deal.